AMENDMENT UNDER 37 C.F.R. § 1.111
U.S. Application No.: 09/556,349

<u>**AMENDMENTS TO THE CLAIMS**</u>

This listing of claims will replace all prior versions and listings of claims in the

application:

LISTING OF CLAIMS:

1. (*Currently Amended*) A method for summarizing a content of an input video sequence,

~~said input video sequence comprising a plurality of frames, said plurality of frames being~~

~~grouped into a plurality of video segments,~~ said method comprising:

(a)   computing a feature vector for each frame in a set of frames from ~~selecting a~~

~~frame cluster in~~ said input video sequence ~~which corresponds to a most static one of said video~~

~~segments;~~ and

(b)   applying singular value decomposition to a matrix comprised of said feature

vectors and projecting the matrix on a refined feature space representation, wherein positions of

said projections on said refined feature space representation represent approximations of visual

changes in said set of frames from said input video sequence~~computing a content value in said~~

~~selected frame cluster;~~

~~(c)   using said computed content value to cluster remaining frames in said input video~~

~~sequence.~~

2-6. (*Cancelled*).

2

7. (*Currently Amended*) The method of claim 1[[6]], wherein said singular value decomposition is performed using frames selected with a fixed interval from said input video sequence.

8. (*Cancelled*).

9. (*Currently Amended*) The method of claim 1[[6]], wherein ~~said singular value decomposition produces a matrix,~~ each column of said matrix represents ~~representing~~ a frame in said ~~a~~ refined feature space representation ~~corresponding to a frame in said input video sequence~~.

10-30. (*Cancelled*).

31. (*Currently Amended*) A computer-readable medium containing a program for summarizing a content of an input video sequence, ~~said input video sequence comprising a plurality of frames, said plurality of frames being grouped into a plurality of video segments,~~ said program comprising:

(a)   computing a feature vector for each frame in a set of frames from ~~selecting a frame cluster in~~ said input video sequence ~~which corresponds to a most static video segment~~; and

(b)   applying singular value decomposition to a matrix comprised of said feature vectors and projecting the matrix on a refined feature space representation, wherein positions of

said projections on said refined feature space representation represent approximations of visual

changes in said set of frames from said input video sequencecomputing content value in said

selected frame cluster;

(c)   using said computed content value to cluster remaining frames in said input video

sequence.

32-36. (*Cancelled*).

37. (*Currently Amended*) The computer-readable medium of claim 31[[36]], wherein said

singular value decomposition is performed using frames selected with a fixed interval from said

input video sequence.

38. (*Cancelled*).

39. (*Currently Amended*) The computer-readable medium of claim 31[[36]], wherein said

singular value decomposition produces a matrix, each column of said matrix represents

representing a frame in said a refined feature space representationcorresponding to a frame in

said input video sequence.

40-72. (*Cancelled*).

73. (*New*) The method of claim 1, wherein said feature vectors are computed using a color histogram that outputs a histogram vector.

74. (*New*) The method of claim 73, wherein said histogram vector is indicative of a spatial distribution of colors in said each of said frames.

75. (*New*) The method of claim 74, wherein each of said frames is divided into a plurality of blocks, each of said plurality of blocks being represented by a histogram in a color space indicative of a distribution of colors within each of said blocks.

76. (*New*) The method of claim 74, wherein each of said frames is divided into a plurality of blocks and said histogram vector comprises a plurality of histograms in a color space, each of said plurality of histograms corresponding to one of said plurality of blocks.

77. (*New*) The method of claim 1, wherein said method further comprises clustering said frames of said input video sequence based upon positions of said projections on said refined feature space representation.

78. (*New*) The method of claim 77, wherein said method further comprises selecting a frame from each cluster to serve as a keyframe in a summarization of said input video sequence.

79. (*New*) The method of claim 78, wherein said selecting a frame comprises locating a

frame with a feature vector that projects into a singular value that is most representative of other

singular values of the cluster.

80. (*New*) The method of claim 78, wherein said method further comprises using said

clustered frames to output a motion video representative of a summary of said input video

sequence.

81. (*New*) The method of claim 80, wherein said input video sequence summary is

composed according to a time-length parameter T_{len} and a minimum display time parameter T_{min}.

82. (*New*) The method of claim 81, wherein the composition of said input video sequence

summary further comprises:

locating the video shot Θ_i in each cluster S_i having the greatest length;

determining how the video shots in each cluster will be arranged according to

$$C \leq N = T_{len}/T_{min},$$

wherein C represents a number of clusters; and

wherein N represents the maximum number of video shots;

if $C \leq N$, then all the video shots in each cluster is included in said input video sequence

summary; and

if $C > N$, then sort each video shot Θ_i from each cluster S_i in descending order by length,

select the first N video shots for inclusion in said input video sequence summary and assign time

length T_{min} to each selected video shot.

83. (*New*) The method of claim 82, wherein the composition of said input video sequence

summary further comprises sorting the selected video shots by their respective time codes.

84. (*New*) The method of claim 83, wherein the composition of said input video sequence

summary further comprises extracting a portion of selected video shot equal in length to time

length T_{min} and inserting each extracted portion in order to said input video sequence summary.

85. (*New*) The method of claim 77, wherein said clustering of said frames further

comprises using a position of the most static shot of said input video sequence to compute a

value as a threshold during the clustering of said frames.

86. (*New*) The method of claim 85, wherein said clustering of said frames further

comprises computing a content value and using said computed content value to cluster the

remaining frames by:

sorting said feature vectors in said refined feature space representation in ascending order

according to a distance of each of said feature vectors to an origin of said refined feature space

representation;

selecting a vector among said sorted feature vectors which is closest to an origin of said

refined feature space representation and including said selected feature vector into a first cluster;

clustering said plurality of sorted feature vectors in said refined feature space

representation into a plurality of clusters according to a distance between each of said plurality of

sorted feature vectors and feature vectors in each of said plurality of clusters and an amount of

information in each of said plurality of clusters.

87. (*New*) The method of claim 86, wherein, in said clustering of sorted feature vectors,

said plurality of sorted feature vectors are clustered into said plurality of clusters such that said

amount of information in each of said plurality of clusters does not exceed an amount of

information in said first cluster.

88. (*New*) The method of claim 86, wherein said first cluster is composed of frames based

on a distance variation between said frames and an average distance between frames in said first

cluster.

89. (*New*) The method of claim 86, wherein each of said plurality of clusters is composed

of frames based on a distance variation between said frames and an average distance between

frames in said each of said plurality of clusters.

90. (*New*) The computer-readable medium of claim 31, wherein said feature vectors are computed using a color histogram that outputs a histogram vector.

91. (*New*) The computer-readable medium of claim 90, wherein said histogram vector is indicative of a spatial distribution of colors in said each of said frames.

92. (*New*) The computer-readable medium of claim 91, wherein each of said frames is divided into a plurality of blocks, each of said plurality of blocks being represented by a histogram in a color space indicative of a distribution of colors within each of said blocks.

93. (*New*) The computer-readable medium of claim 91, wherein each of said frames is divided into a plurality of blocks and said histogram vector comprises a plurality of histograms in a color space, each of said plurality of histograms corresponding to one of said plurality of blocks.

94. (*New*) The computer-readable medium of claim 31, wherein said method further comprises clustering said frames of said input video sequence based upon positions of said projections on said refined feature space representation.

9

95. (*New*) The computer-readable medium of claim 94, wherein said method further comprises selecting a frame from each cluster to serve as a keyframe in a summarization of said input video sequence.

96. (*New*) The computer-readable medium of claim 95, wherein said selecting a frame comprises locating a frame with a feature vector that projects into a singular value that is most representative of other singular values of the cluster.

97. (*New*) The computer-readable medium of claim 95, wherein said method further comprises using said clustered frames to output a motion video representative of a summary of said input video sequence.

98. (*New*) The computer-readable medium of claim 97, wherein said input video sequence summary is composed according to a time-length parameter T_{len} and a minimum display time parameter T_{min}.

99. (*New*) The computer-readable medium of claim 98, wherein the composition of said input video sequence summary further comprises:

locating the video shot Θ_i in each cluster S_i having the greatest length;

determining how the video shots in each cluster will be arranged according to

$$C \leq N = T_{len}/T_{min},$$

wherein C represents a number of clusters; and

wherein N represents the maximum number of video shots;

if $C \leq N$, then all the video shots in each cluster is included in said input video sequence summary; and

if $C > N$, then sort each video shot Θ_i from each cluster S_i in descending order by length, select the first N video shots for inclusion in said input video sequence summary and assign time length T_{min} to each selected video shot.

100. (*New*) The computer-readable medium of claim 99, wherein the composition of said input video sequence summary further comprises sorting the selected video shots by their respective time codes.

101. (*New*) The computer-readable medium of claim 100, wherein the composition of said input video sequence summary further comprises extracting a portion of selected video shot equal in length to time length T_{min} and inserting each extracted portion in order to said input video sequence summary.

102. (*New*) The computer-readable medium of claim 94, wherein said clustering of said frames further comprises using a position of the most static shot of said input video sequence to compute a value as a threshold during the clustering of said frames.

103. (*New*) The computer-readable medium of claim 102, wherein said clustering of said frames further comprises computing a content value and using said computed content value to cluster the remaining frames by:

sorting said feature vectors in said refined feature space representation in ascending order according to a distance of each of said feature vectors to an origin of said refined feature space representation;

selecting a vector among said sorted feature vectors which is closest to an origin of said refined feature space representation and including said selected feature vector into a first cluster;

clustering said plurality of sorted feature vectors in said refined feature space representation into a plurality of clusters according to a distance between each of said plurality of sorted feature vectors and feature vectors in each of said plurality of clusters and an amount of information in each of said plurality of clusters.

104. (*New*) The computer-readable medium of claim 103, wherein, in said clustering of sorted feature vectors, said plurality of sorted feature vectors are clustered into said plurality of clusters such that said amount of information in each of said plurality of clusters does not exceed an amount of information in said first cluster.

105. (*New*) The computer-readable medium of claim 103, wherein said first cluster is composed of frames based on a distance variation between said frames and an average distance between frames in said first cluster.

106. (*New*) The computer-readable medium of claim 103, wherein each of said plurality of clusters is composed of frames based on a distance variation between said frames and an average distance between frames in said each of said plurality of clusters.